41076: Methods in Quantum Computing

'Entanglement Transformation' Module

Min-Hsiu Hsieh
Centre for Quantum Software & Information, Faculty of Engineering and Information Technology, University of Technology Sydney

Abstract

Contents to be covered in this lecture are

1. Condition for exact entanglement transformation
2. Asymptotical reversible?
3. Quantumness and Total Correlation
4. State Merging

In this module, we will introduce fundamental results on entanglement transformations. These are the most significant results in entanglement theory in my opinion. Note that I look at the dynamic picture of entanglement instead of the static properties (such as bell inequality, entanglement structures etc). What I mean by dynamic picture is how transformation of entanglement helps us to better understand the resource of entanglement.

Preliminary

Without loss of generality, we always assume a bipartite pure states $|\phi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ has the Schmidt decomposition in the standard basis:

$$|\phi\rangle_{AB} = \sum_{i=1}^{d} \sqrt{p_i}|i\rangle_A \otimes |i\rangle_B. \tag{1}$$

Otherwise, there exist local unitary operations that transform the state to the form in Eq. (1). Furthermore, we denote $|\phi\rangle_{AB} \sim |\varphi\rangle_{AB}$ if two bipartite pure states $|\phi\rangle_{AB}$ and $|\varphi\rangle_{AB}$ are equivalent up to local unitary operations. Denote the reduced density matrix of $|\phi\rangle_{AB}$ in Eq. (1) by

$$\rho_\phi := \mathrm{Tr}_A|\phi\rangle\langle\phi| = \mathrm{Tr}_B|\phi\rangle\langle\phi| = \sum_{i=1}^{d} p_i|i\rangle\langle i|. \tag{2}$$

It is also easy to check that the eigenvalues of ρ_ϕ and ρ_φ are the same if $|\phi\rangle_{AB} \sim |\varphi\rangle_{AB}$.

Majorization

Consider two d-dimension real vectors $\boldsymbol{x} = (x_1, \cdots, x_d)$ and $\boldsymbol{y} = (y_1, \cdots, y_d)$. We say $\boldsymbol{x}$ is majorized by $\boldsymbol{y}$, denoted by $\boldsymbol{x} \prec \boldsymbol{y}$, if for all $k \in [d]$,

$$\sum_{j=1}^{k} x_j^{\downarrow} \leq \sum_{j=1}^{k} y_j^{\downarrow}, \tag{3}$$

where the equality holds when $k = d$, and $x_j^{\downarrow}$ denotes the j-th largest element with $x_1^{\downarrow}$ being the largest element in $\boldsymbol{x}$.

Lemma 1. *If $\boldsymbol{x} \prec \boldsymbol{y}$, then*

- *there exists a double stochastic matrix D, where $\sum_{i=1}^{d} D_{i,j} = \sum_{j=1}^{d} D_{i,j} = 1$, such that*

$$\boldsymbol{x} = D\boldsymbol{y}.$$

- *for any concave function f*

$$\sum_{j=1}^{d} f(x_j) \leq \sum_{j=1}^{d} f(y_j). \tag{4}$$

Entropy

For a quantum state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ with

$$\rho_A = \sum_{i=1}^{d} p_i |i\rangle\langle i|_A,$$

its von Neumann entropy

$$H(A)_\rho := H(\rho_A) = -\operatorname{Tr} \rho_A \log \rho_A \tag{5}$$

is also equal to the Shannon entropy $H(p)$ defined on the probability distribution p. The entanglement entropy $E(\psi)$ of a bipartite pure state $|\psi\rangle_{AB}$ is equal to $H(\rho_\psi)$.

The conditional von Neumann entropy $H(A|B)_\sigma$ of $\sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ is define as

$$H(A|B)_\sigma = H(A)_\sigma - H(AB)_\sigma. \tag{6}$$

Exercise 2. *Show that $H(X|Y)$ of a joint distribution p_{XY} is always non-negative.*

Exercise 3. *Show by example that $H_(A|B)_\sigma$ could be negative. Why is this?*

The quantum mutual information $I(A : B)_\sigma$ is defined as

$$I(A : B)_\sigma = H(A)_\sigma + H(B)_\sigma - H(AB)_\sigma. \tag{7}$$

Exercise 4. *Find out the condition for when the quantum mutual information $I(A : B)_\sigma = 0$.*

The quantum mutual information is continuous in the sense that if two states are close in trace distance, then their quantum mutual informations will also be close. The following lemma makes this intuition quantitative.

Lemma 5. *It holds*

$$\|I(A:B)_\rho - I(A:B)_\sigma| \leq 6\epsilon \log d_A + 4H_2(\epsilon) \tag{8}$$

for any ρ_{AB} and σ_{AB} satisfying

$$\|\rho_{AB} - \sigma_{AB}\|_1 \leq \epsilon. \tag{9}$$

Here

$$H_2(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon).$$

Lemma 6. *Let $\sigma_{AB} = I_B \otimes \mathcal{N}_A(\rho_{AB})$. We have*

$$I(A:B)_\rho \geq I(A:B)_\sigma. \tag{10}$$

For two quantum states ρ and σ, we define the relative entropy

$$D(\rho\|\sigma) = \operatorname{Tr} \rho(\log \rho - \log \sigma). \tag{11}$$

The relative entropy is an important quantity for many applications, and has many nice properties

1. Monotonicity:
$$D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq D(\rho\|\sigma).$$

2. Joint Convexity
$$D(\lambda\rho_1 + (1-\lambda)\rho_2\|\lambda\sigma_1 + (1-\lambda)\sigma_2) \leq \lambda D(\rho_1\|\sigma_1) + (1-\lambda)D(\rho_2\|\sigma_2).$$

Exercise 7. *Show that $D(\rho_{AB}\|\rho_A \otimes \rho_B) = I(A:B)_\rho$.*

1 Condition for exact entanglement transformation

The first fundamental result here is the condition for when a pure entangled state $|\psi\rangle_{AB}$ can be transformed to $|\phi\rangle$ [6] if only local operations and unlimited two-way classical communication (LOCC) are allowed. Asher Peres and William Wootters were the first to introduce the LOCC paradigm and study it as a restricted class of operations in their seminal work [8].

In the LOCC protocol, each party is allowed to perform the following operations:

(i) Perform a local quantum instrument $(\mathcal{E}_m)_m$ [3], where each $\mathcal{E}_m$ is a completely positive (CP) map, and their sum $\sum_m \mathcal{E}_m$ is a trace-preserving map. Quantum instruments represent the most general type of quantum measurement. When performing the instrument on the state σ, the "measurement" outcome m is obtained with probability $p(m) = tr[\mathcal{E}_m(\sigma)]$, and the post-measurement state given this outcome is $\sigma_m = \mathcal{E}_m(\sigma)/p(m)$ for $p(m) > 0$.

(ii) Broadcast the result of any quantum measurement.

A general LOCC protocol is described by a multi-level "tree" of local instruments in which the choice of instrument performed at each node of the tree depends on the particular history of measurement outcomes leading up to that node.

Before proving the general result, let's first look at a special example as follows.

Lemma 8. *Given $|\psi\rangle_{AB}, |\phi\rangle_{AB} \in \mathcal{H}_2 \otimes \mathcal{H}_2$,*

$$
\begin{aligned}
|\psi\rangle_{AB} &= \sqrt{\alpha_0}|0\rangle_A|0\rangle_B + \sqrt{\alpha_1}|1\rangle_A|1\rangle_B & (12)\\
|\phi\rangle_{AB} &= \sqrt{\beta_0}|0\rangle_A|0\rangle_B + \sqrt{\beta_1}|1\rangle_A|1\rangle_B & (13)
\end{aligned}
$$

where $\beta_1 \leq \alpha_1 \leq \alpha_0 \leq \beta_0$ and $\sum_{i=0}^{1} \alpha_i = \sum_{i=0}^{1} \beta_i = 1$. Construct a LOCC protocol to transform $|\psi\rangle_{AB}$ to $|\phi\rangle_{AB}$.

Proof. In the following, we will show that a one-way protocol suffices to transform $|\psi\rangle_{AB}$ to $|\phi\rangle_{AB}$.

Notice that

$$
|\psi\rangle_{AB} \sim |\psi'\rangle_{AB} = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A(\cos\gamma|0\rangle_B + \sin\gamma|1\rangle_B)) \tag{14}
$$

if $\alpha_0 = \frac{1+\cos\gamma}{2}$. Next, Alice can perform a two-outcome measurement with the following measurement operators:

$$
M_0 = \begin{pmatrix} \cos\delta & 0 \\ 0 & \sin\delta \end{pmatrix}, \quad M_1 = \begin{pmatrix} \sin\delta & 0 \\ 0 & \cos\delta \end{pmatrix}. \tag{15}
$$

Denote $|\psi_0''\rangle_{AB}$ and $|\psi_1''\rangle_{AB}$ be the residue states after observing outcome '0' and '1', respectively. A simple calculation shows that

$$
\begin{aligned}
|\psi_0''\rangle_{AB} &= \cos\delta|00\rangle_{AB} + \sin\delta|1\rangle_A(\cos\gamma|0\rangle_B + \sin\gamma|1\rangle_B) & (16)\\
|\psi_1''\rangle_{AB} &= \sin\delta|00\rangle_{AB} + \cos\delta|1\rangle_A(\cos\gamma|0\rangle_B + \sin\gamma|1\rangle_B). & (17)
\end{aligned}
$$

Moreover, $|\psi_0''\rangle_{AB} \sim |\psi_1''\rangle_{AB}$. Let

$$
|\psi''\rangle_{AB} = \sqrt{\lambda_+}|00\rangle_{AB} + \sqrt{\lambda_-}|11\rangle_{AB}, \tag{18}
$$

where

$$
\lambda_\pm = \frac{1 \pm \sqrt{1 - \sin^2(2\delta)\sin^2(\gamma)}}{2}. \tag{19}
$$

We can verify that $|\psi''\rangle_{AB} \sim |\psi_0''\rangle_{AB} \sim |\psi_1''\rangle_{AB}$. Since $\alpha_0 \leq \beta_0 \leq 1$, $\lambda_+(\delta = 0) = 1$ and $\lambda_+(\delta = \pi/4) = \alpha_0$, choosing $\delta \in [0, \pi/4]$

$$
\delta = \frac{1}{2}\arcsin\left(2\sqrt{\beta_0 - \beta_0^2}/\sin\gamma\right) \tag{20}
$$

we will have $|\psi''\rangle \sim |\phi\rangle$.

Note that

$$
\rho_\psi = \begin{pmatrix} \alpha_0 & 0 \\ 0 & \alpha_1 \end{pmatrix}, \quad \rho_\phi = \begin{pmatrix} \beta_0 & 0 \\ 0 & \beta_1 \end{pmatrix}. \tag{21}
$$

The above LOCC protocol is equivalent to perform a double stochastic transformation

$$
T = \begin{pmatrix} t & 1-t \\ 1-t & t \end{pmatrix}. \tag{22}
$$

Namely, $\lambda_\psi = T\lambda_\phi$, when

$$
t = \frac{\alpha_0 - \beta_1}{\beta_0 - \beta_1}.
$$

□

Now we are ready to state the main result of exact entanglement transformation under LOCC.

Theorem 9. $|\psi\rangle_{AB} \xrightarrow{LOCC} |\phi\rangle_{AB}$ *if and only if* $\lambda_\psi \prec \lambda_\phi$, *where* λ_ψ *and* λ_ϕ *are Schmidt coefficients of* $|\psi\rangle_{AB}$ *and* $|\phi\rangle_{AB}$, *respectively.*

$\rightarrow$. We only need to consider the one-way protocol. Suppose that the transformation from $|\psi\rangle_{AB}$ to $|\phi\rangle_{AB}$ can be done by Alice's local measurement $\{M_i\}$, followed by Bob's local operation $\mathcal{E}_i$. In other words, for every i,

$$|\phi\rangle\langle\phi| \approx (I_A \otimes \mathcal{E}_i)((M_i \otimes I_B)|\psi\rangle\langle\psi|(M_i^\dagger \otimes I_B)), \tag{23}$$

otherwise, Bob's output state cannot be a pure state. Here we omit the normalization factor due to the quantum measurement. Partially tracing out subsystem B, we obtain

$$\rho_\phi = \mathrm{Tr}_B|\phi\rangle\langle\phi| = \frac{M_i\rho_\psi M_i^\dagger}{p_i}. \tag{24}$$

Let the polar decomposition of $A \equiv M_i\sqrt{\rho_\psi}$ be $|A|U_i$ for some unitary U_i, where $|A| = \sqrt{AA^\dagger}$:

$$M_i\sqrt{\rho_\psi} = \sqrt{M_i\rho_\psi M_i^\dagger}U_i = \sqrt{p_i\rho_\phi}U_i \tag{25}$$

Because we can rewrite

$$\rho_\psi = \sum_i \sqrt{\rho_\psi}M_i^\dagger M_i\sqrt{\rho_\psi} \tag{26}$$

$$= \sum_i p_i U_i^\dagger \rho_\phi U_i \tag{27}$$

where the last equality holds only if $\lambda_\psi \prec \lambda_\phi$. □

$\leftarrow$. We employ Lemma 1 and let D be the double stochastic matrix so that $\rho_\psi = D\rho_\phi$. Note that we can decompose $D = \prod_k T_k$, where for each k, T_k is a 2×2 matrix of the form

$$T_k = \begin{pmatrix} t & 1-t \\ 1-t & t \end{pmatrix}. \tag{28}$$

The idea here is to construct sequential entanglement transformations, each only acts on a two-dimensional space. The effect of these entanglement transformations will equivalently equal to the sequential T transforms transferring ρ_ϕ to ρ_ψ. Without loss of generality, we assume that the Schmidt decomposition of $|\psi\rangle$ and $|\phi\rangle$:

$$|\psi\rangle = \sum_{i=0}^{d-1} \sqrt{\alpha_i}|i\rangle|i\rangle \tag{29}$$

$$|\phi\rangle = \sum_{i=0}^{d-1} \sqrt{\beta_i}|i\rangle|i\rangle. \tag{30}$$

To do that we first notice that

$$|\psi\rangle \sim |\psi'\rangle = \cos\xi(\sqrt{\alpha_0}|00\rangle + \sqrt{\alpha_1}|11\rangle) + \sin\xi|\psi_\perp\rangle \tag{31}$$

where $|\psi_\perp\rangle \approx \sum_{i=2}^{d-1} \alpha_i |i\rangle|i\rangle$. Similarly,

$$|\phi\rangle \sim |\phi'\rangle = \cos\xi\Big(\sqrt{\beta_0}|00\rangle + \sqrt{\beta_1}|11\rangle\Big) + \sin\xi|\psi_\perp\rangle \tag{32}$$

where $|\psi_\perp\rangle \approx \sum_{i=2}^{d-1} \alpha_i |i\rangle|i\rangle$. The measurement that Alice performs is

$$\tilde{M}_0 = \begin{pmatrix} M_0 & 0 \\ 0 & \frac{1}{\sqrt{2}} I_{d-2} \end{pmatrix}, \quad \tilde{M}_1 = \begin{pmatrix} M_1 & 0 \\ 0 & \frac{1}{\sqrt{2}} I_{d-2} \end{pmatrix} \tag{33}$$

where M_0, M_1 are given in Eq. (15).

□

2 Asymptotically reversible

The entanglement entropy is a measure of quantum entanglement in a bipartite state. It is easy to see that entanglement entropy is zero for separable states and is equal to one for maximally entangled states. The entanglement entropy has the following operational meanings. For a given bipartite pure state $|\psi\rangle$, we can achieve

$$|\psi\rangle_{AB}^{\otimes n} \xleftrightarrow{\text{LOCC}} |\Phi_+\rangle^{\otimes nE(\psi)}, \quad n \to \infty \tag{34}$$

where $|\Phi_+\rangle$ is the maximally entangled state

$$|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}).$$

The $\to$ direction is the entanglement concentration protocol; while the $\leftarrow$ direction is called the entanglement dilution protocol. Therefore, a bipartite pure entangled state is asymptotically reversible. However, this is not true for mixed entangled states because of the existence of bound entangled states. If you are an experimentalist, this is all you need to know. In the following subsections, I will detail the proofs for Eq. (34).

2.1 Entanglement Concentration

Denote $|\psi\rangle_{AB}^{\otimes n} := |\psi\rangle_{A^n B^n}$ and denote the maximally entangled state of size N by

$$|\Phi_N\rangle_{\hat{A}\hat{B}} = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle_{\hat{A}} \otimes |i\rangle_{\hat{B}}.$$

We formally define the protocol of entanglement concentration as follows (see Figure 1).

Definition 10 (Entanglement Concentration)**.** *An $(n, E-\delta, \epsilon)$ entanglement concentration protocol to transform n copies of $|\psi\rangle_{AB}$ into the maximally entangled state $|\Phi_N\rangle_{\hat{A}\hat{B}}$ of size N, where $\frac{1}{n}\log_2 N = E - \delta$, consists of a pair of operations $\mathcal{E} : A^n \to \hat{A}$ and $\mathcal{F} : B^n \to \hat{B}$ by Alice and Bob, respectively, so that the resulting state*

$$\omega_{\hat{A}\hat{B}} = \Big(\mathcal{E}^{A^n\to\hat{A}} \otimes \mathcal{F}^{B^n\to\hat{B}}\Big)(|\psi\rangle\langle\psi|_{A^nB^n}) \tag{35}$$

is ϵ-close to $|\Phi_N\rangle_{\hat{A}\hat{B}}$:

$$\frac{1}{2}\big\|\omega_{\hat{A}\hat{B}} - |\Phi_N\rangle\langle\Phi_N|_{\hat{A}\hat{B}}\big\|_1 \le \epsilon. \tag{36}$$

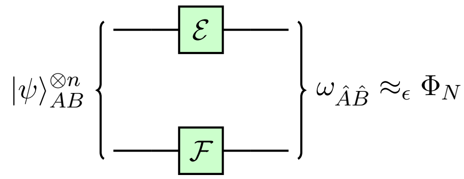

Figure 1: The entanglement concentration protocol.

We say that the entanglement concentration rate E is *achievable* if there exists an $(n, E-\delta, \epsilon)$ entanglement concentration protocol for all $\delta, \epsilon > 0$ and sufficiently large n.

Theorem 11 ([1])**.**

$$\sup\{E : E \textit{ is achievable}\} = E(\psi). \tag{37}$$

The information theoretical approach to prove Theorem 11 consists of two parts: the converse proof and the direct coding theorem. The converse proof argues that any entanglement concentration protocol that is ϵ-error must has the rate $\frac{1}{n}\log N$ less than the entanglement entropy $E(\psi)$. On the other hand, the direct coding theorem explicitly constructs local quantum operations for Alice and Bob to distill entanglement of size $N \approx nE(\psi)$.

The converse proof

For any entanglement concentration protocol that achieves Eq. (36), it will satisfy

$$\begin{aligned}
\log N &= H(\hat{A})_{\Phi_N} && (38)\\
&= \frac{1}{2} I(\hat{A} : \hat{B})_{\Phi_N} && (39)\\
&\le \frac{1}{2} I(\hat{A} : \hat{B})_{\omega} + f(n,\epsilon) && (40)\\
&\le \frac{1}{2} I(A^n : B^n)_{\psi_n} + f(n,\epsilon) && (41)\\
&= H(A^n)_{\psi_n} + f(n,\epsilon) && (42)\\
&= nE(\psi) + f(n,\epsilon). && (43)
\end{aligned}$$

The first inequality follows from Lemma 5 with $f(n,\epsilon) = 6\epsilon \log d_A + 4H_2(\epsilon)$. The second inequality follows from the data processing inequality in Lemma 6. The state $\psi_n := |\psi\rangle_{AB}^{\otimes n}$.

The direct coding theorem

We will need a common tool to prove the direct coding theorem. The method of types is a standard technique of classical information theory. Denote by $\boldsymbol{x}$ a sequence $x_1 x_2 \ldots x_n$, where each x_i belongs to the finite set $\mathcal{X}$. Denote by $|\mathcal{X}|$ the cardinality of $\mathcal{X}$. Denote by $N(a|\boldsymbol{x})$ the number of occurrences of the symbol $a \in \mathcal{X}$ in the sequence $\boldsymbol{x}$. The *type* $t^{\boldsymbol{x}}$ of a sequence $\boldsymbol{x}$ is a probability vector:

$$t^{\boldsymbol{x}} = \begin{pmatrix} t_1^{\boldsymbol{x}} \\ t_2^{\boldsymbol{x}} \\ \vdots \\ t_{|\mathcal{X}|}^{\boldsymbol{x}} \end{pmatrix}, \tag{44}$$

where the a-th entry is $t^{\boldsymbol{x}}_a = \frac{N(a|\boldsymbol{x})}{n}$. Denote the set of sequences of type t by

$$\mathcal{T}^n_t = \{\boldsymbol{x} \in \mathcal{X}^n : t^{\boldsymbol{x}} = t\}.$$

For the probability distribution p on the set $\mathcal{X}$ and $\delta > 0$, let $\tau_\delta = \{t : \forall a \in \mathcal{X},\ |t_a - p_a| \le \delta\}$. Define the set of δ-typical sequences of length n as

$$\begin{aligned} \mathcal{T}^n_{p,\delta} &= \bigcup_{t\in\tau_\delta} \mathcal{T}^n_t \\ &= \{\boldsymbol{x} : \forall a \in \mathcal{X},\ |t^{\boldsymbol{x}}_a - p_a| \le \delta\}. \end{aligned} \tag{45}$$

Define the probability distribution p^n on $\mathcal{X}^n$ to be the tensor power of p. The sequence $\boldsymbol{x}$ is drawn from p^n if and only if each letter x_i is drawn independently from p. Typical sequences enjoy many useful properties. Let $H(p) = -\sum_x p_x \log p_x$ be the Shannon entropy of p. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$p^n(\mathcal{T}^n_{p,\delta}) \ge 1 - \epsilon \tag{46}$$

$$2^{-n[H(p)+c\delta]} \le p^n(\boldsymbol{x}) \le 2^{-n[H(p)-c\delta]},\ \ \forall \boldsymbol{x} \in \mathcal{T}^n_{p,\delta} \tag{47}$$

$$|\mathcal{T}^n_{p,\delta}| \le 2^{n[H(p)+c\delta]} \tag{48}$$

for some constant c (see [2] for proofs). For $t \in \tau_\delta$ and for sufficiently large n, the cardinality $D_t = |\mathcal{T}^n_t|$ is bounded as [2]

$$2^{n[H(p)+\eta(\delta)]} \ge D_t \ge 2^{n[H(p)-\eta(\delta)]} \tag{49}$$

where

$$n\eta(\delta) = B\delta\sqrt{n} + |\mathcal{X}| \log_2(n+1) + C\delta^2 \tag{50}$$

for some constant B and C. Note that the function $\eta(\delta) \to 0$ as $n \to \infty$.

The above concepts generalize to the quantum setting by virtue of the spectral theorem. Let $\rho = \sum_{x\in\mathcal{X}} p_x |x\rangle\langle x|$ be the spectral decomposition of a given density matrix ρ. In other words, $|x\rangle$ is the eigenstate of ρ corresponding to eigenvalue p_x. The von Neumann entropy of the density matrix ρ is

$$H(\rho) = -\operatorname{Tr} \rho \log \rho = H(p).$$

Define the type projector

$$\Pi^n_t = \sum_{\boldsymbol{x}\in\mathcal{T}^n_t} |\boldsymbol{x}\rangle\langle\boldsymbol{x}|. \tag{51}$$

The density operator proportional to the type projector is $\pi_t = D_t{}^{-1}\Pi^n_t$. The typical subspace associated with the density matrix ρ is defined as

$$\Pi^n_{\rho,\delta} = \sum_{\boldsymbol{x}\in\mathcal{T}^n_{p,\delta}} |\boldsymbol{x}\rangle\langle\boldsymbol{x}| = \sum_{t\in\tau_\delta} \Pi^n_t. \tag{52}$$

Properties analogous to (46) – (49) hold [7]. For any $\epsilon, \delta > 0$, and all sufficiently large n for which

$$\operatorname{Tr} \rho^{\otimes n} \Pi^n_{\rho,\delta} \ge 1 - \epsilon \tag{53}$$

$$2^{-n[H(\rho)+c\delta]}\Pi^n_{\rho,\delta} \le \Pi^n_{\rho,\delta}\rho^{\otimes n}\Pi^n_{\rho,\delta} \le 2^{-n[H(\rho)-c\delta]}\Pi^n_{\rho,\delta}, \tag{54}$$

$$\mathrm{Tr}\,\Pi^n_{\rho,\delta} \le 2^{n[H(\rho)+c\delta]} \tag{55}$$

for some constant c. For $t \in \tau_\delta$ and for sufficiently large n, the dimension of the type projector Π^n_t is lower bounded as

$$2^{n[H(\rho)+\eta(\delta)]} \ge \mathrm{Tr}\,\Pi^n_t \ge 2^{n[H(\rho)-\eta(\delta)]} \tag{56}$$

and the function $\eta(\delta)$ is given in Eq. (50).

Direct Coding Theorem. Let $\rho = \mathrm{Tr}|\psi\rangle\langle\psi|_{AB}$. Let $\Pi^n_{\rho,\delta}$ and Π^n_t be the typical and type projectors defined in Eqs. (52) and (51), respectively. Let $A_n \equiv \mathcal{H}_A^{\otimes n}$. Let $\hat{A}_n \subset A_n$ be the subspace spanned by $\Pi^n_{\rho,\delta}$ and let $\tilde{A}_n = A_n \backslash \hat{A}_n$. Similarly, we define $B_n, \tilde{B}_n, \hat{B}_n$.

Denote $\Upsilon_0 = I_{A_n} - \Pi^n_{\rho,\delta}$, $\Upsilon_t = \Pi^n_t \Pi^n_{\rho,\delta}$, $\forall t \in [\tau_\delta]$. Define $\mathcal{E}_A : A_n \to YA_n$ by

$$\mathcal{E}_A(\sigma_{A_n}) = \sum_{t=0}^{|\tau_\delta|} |t\rangle\langle t|_Y \otimes \Upsilon_t(\sigma_{A_n})\Upsilon_t. \tag{57}$$

Exercise 12. *Verify that $\mathcal{E}_A$ in Eq. (57) is a CPTP map.*

We can define $\mathcal{F}_B : B_n \to YB_n$ in the exact same way:

$$\mathcal{F}_B(\sigma_{B_n}) = \sum_{t=0}^{|\tau_\delta|} |t\rangle\langle t|_Y \otimes \Upsilon_t(\sigma_{B_n})\Upsilon_t. \tag{58}$$

The entanglement concentration protocol consists of Alice and Bob performing $\mathcal{E}_A$ and $\mathcal{F}_B$, followed by measuring the Y register. For any measurement outcome $t \neq 0$, the protocol succeeds and a maximally entangled state $|\Phi_t\rangle$ will be created

$$|\Phi_t\rangle_{A_nB_n} = \frac{1}{\sqrt{D_t}} \sum_{\boldsymbol{x} \in \mathcal{T}^n_t} |\boldsymbol{x}\rangle_{A_n} \otimes |\boldsymbol{x}\rangle_{B_n}. \tag{59}$$

It follows from Eq. (49) that

$$\frac{1}{n}\log_2 D_t \ge H(\rho) - \eta(\delta). \tag{60}$$

Therefore the rate of entanglement concentration protocol goes to $E(\psi) = H(\rho)$ when n goes to infinity because $\eta(\delta) \to 0$ ($\delta \to 0$).

What remains to show is that the error of the aforementioned entanglement concentration protocol is small. Note that the only instance that the protocol would fail is when the measurement outcome $t = 0$. Note that the probability that this would happen is

$$\Pr\{t = 0\} = \mathrm{Tr}\,\Upsilon_0 \rho^{\otimes n} \le \epsilon, \tag{61}$$

where the inequality follows from Eq. (53).

Exercise 13. *Please made the error analysis more rigorous!*

□

2.2 Entanglement Dilution

The reverse protocol of entanglement concentration is called entanglement dilution. Given the maximally entangled state

$$|\Phi_N\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle_{\hat{A}} \otimes |i\rangle_{\hat{B}},$$

the goal is to obtain $|\psi\rangle$.

We formally define the protocol of entanglement dilution as follows.

Definition 14 (Entanglement Dilution)**.** *An $(n, E+\delta, \epsilon)$ entanglement dilution protocol to transform the preshared maximally entangled state $|\Phi_N\rangle_{\hat{A}\hat{B}}$ of size N, where $\frac{1}{n}\log_2 N = E + \delta$, into n copies of $|\psi\rangle_{AB}$ consists*

- *Alice's instrument $\{\mathcal{E}_x : \hat{A} \to A^n X\}_{x\in\mathcal{X}}$ and send classical index x to Bob;*
- *Bob's decoding operation $\{U_x : \hat{B} \to B^n\}_{x\in\mathcal{X}}$*

so that

$$\omega_{A^nB^nX} = \sum_{x\in\mathcal{X}} (\mathcal{E}_x \otimes U_x)(|\Phi_N\rangle\langle\Phi_N|) \tag{62}$$

is ϵ-close to $|\psi\rangle_{A^nB^n}$:

$$\frac{1}{2}\left\|\omega_{\hat{A}\hat{B}} - |\psi\rangle\langle\psi|_{A^nB^n}\right\|_1 \le \epsilon. \tag{63}$$

We say that the entanglement dilution rate E is *achievable* if there exists an $(n, E+\delta, \epsilon)$ entanglement dilution protocol for all $\delta, \epsilon > 0$ and sufficiently large n.

Theorem 15 ([1,5])**.**

$$\inf\{E : E \text{ is achievable}\} = E(\psi). \tag{64}$$

The converse proof

For any entanglement dilution protocol that achieves Eq. (63), it will satisfy

$$\begin{aligned}
\log N &= H(\hat{A})_{\Phi_N} && (65)\\
&= \frac{1}{2} I(\hat{A}:\hat{B})_{\Phi_N} && (66)\\
&\ge \frac{1}{2} I(\hat{A}:\hat{B})_{\omega} && (67)\\
&\ge \frac{1}{2} I(A^n:B^n)_{\psi_n} - f(n,\epsilon) && (68)\\
&= H(A^n)_{\psi_n} - f(n,\epsilon) && (69)\\
&= nE(\psi) - f(n,\epsilon). && (70)
\end{aligned}$$

The first inequality follows from the data processing inequality in Lemma 6. The second inequality follows from Lemma 5 with $f(n,\epsilon) = 6\epsilon \log d_A + 4H_2(\epsilon)$. The state $\psi_n := |\psi\rangle_{AB}^{\otimes n}$.

The direct coding theorem

Direct Coding Theorem. First of all, we can ignore the atypical subspace since its contribution is very small. Next, we can decompose the typical subspace $\Pi^n_{\rho,\delta}$ into $\Phi_d \otimes \Delta$, where

$$\log_2 d = n[H(\rho) - 2\eta(\delta)]. \tag{71}$$

Note that, $\forall t \in \tau_\delta$, there exists a bijection map $e_t : \mathcal{T}^n_t \to \mathcal{I} \times \mathcal{J}_t$, where $|\mathcal{I}| = d$ and $\mathcal{J}_t = \mathcal{T}^n_t \backslash \mathcal{I}$, so that $e_t(\boldsymbol{x}) = (i, j)$. Then

$$\begin{aligned}
|\Phi_t\rangle &= \frac{1}{\sqrt{D_t}} \sum_{\boldsymbol{x}\in\mathcal{T}^n_t} |\boldsymbol{x}\rangle \otimes |\boldsymbol{x}\rangle & (72)\\
&= \frac{1}{\sqrt{D_t}} \sum_{i\in\mathcal{I}} \sum_{j\in\mathcal{J}_t} |i, j\rangle \otimes |i, j\rangle & (73)\\
&= \left(\frac{1}{\sqrt{d}} \sum_i |i\rangle \otimes |i\rangle\right) \otimes \left(\frac{\sqrt{d}}{\sqrt{D_t}} \sum_j |j\rangle \otimes |j\rangle\right) & (74)\\
&= |\Phi_d\rangle \otimes |\Delta_t\rangle & (75)
\end{aligned}$$

where

$$|\Delta_t\rangle := \frac{\sqrt{d}}{\sqrt{D_t}} \sum_{j\in\mathcal{J}_t} |j\rangle \otimes |j\rangle. \tag{76}$$

It follows from Eq. (49) that

$$\begin{aligned}
\dim(\Delta_t) = |\mathcal{J}_t| &\leq \frac{|\mathcal{T}^n_t|}{d} & (77)\\
&\leq 2^{3n\eta(\delta)}. & (78)
\end{aligned}$$

Moreover,

$$\begin{aligned}
|\psi\rangle_{\tilde{A}_n} &= \frac{1}{|\tau_\delta|} \sum_t |\Phi_t\rangle & (79)\\
&= |\Phi_d\rangle \otimes |\Delta\rangle & (80)
\end{aligned}$$

where

$$|\Delta\rangle := \frac{1}{|\tau_\delta|} \sum_t |\Delta_t\rangle \tag{81}$$

The dimension of $|\Delta\rangle$ is upper bounded by

$$(n+1)^{|\mathcal{X}|} 2^{3n\eta(\delta)} = 2^{3n\eta(\delta)+|\mathcal{X}|\log_2(n+1)}.$$

Let us summarize the entanglement dilation protocol. Assume Alice and Bob initially share maximally entangled state of size N, where $\log_2 N := nH(\rho) + 3n\eta(\delta) + |\mathcal{X}|\log_2(n+1)$. Alice first locally prepares the state $|\Delta\rangle$ in Eq. (81). Then they perform quantum teleportation to send half of $|\Delta\rangle$ to Bob by using this amount of entanglement: $3n\eta(\delta) + |\mathcal{X}|\log_2(n+1)$. Once this is done, Alice and Bob share the state $|\Phi_d\rangle \otimes |\Delta\rangle$ in Eq. (80). It is then quite trivial that

$$\| |\Phi_d\rangle \otimes |\Delta\rangle - |\psi\rangle^{\otimes n}_{AB} \| \leq \epsilon.$$

Finally,

$$\lim_{n\to\infty} \frac{1}{n} \log_2 N = H(\rho). \tag{82}$$

□

2.3 Entanglement is asymptotically reversible.

Combining Theorems 11 and 15, we can see that the entanglement resource in pure bipartite states is asymptotically reversible. However, there is a caveat in this statement. It has been shown that additional classical communication is needed in the task of entanglement dilution. This is not the case for entanglement concentration.

Exercise 16 ([4])**.** *Show that the minimum amount of classical communication in entanglement dilution is $O(\sqrt{n})$.*

Finally, I would like to emphasise again that the entanglement resource in general bipartite mixed states is not reversible.

Exercise 17. *Construct an entangled state that cannot be distilled using any local operations and classical communication (LOCC) procedures. [hint: bound entangled states]*

3 Quantumness versus Total Correlations

In classical information theory, mutual information is used to describe the correlation between joint distribution p_{XY}. We denote

$$I(X:Y) = H(X) + H(Y) - H(XY). \tag{83}$$

The mutual information $I(X:Y) = 0$ if and only if $p_{XY} = p_X p_Y$, i.e., they are independent.

Correspondingly, the quantum mutual information of a bipartite state ρ_{AB} can be defined as

$$I(A:B) = H(A) + H(B) - H(AB) \tag{84}$$

where $H(\cdot)$ is the von Neumann entropy of the system. The quantum mutual information $I(A:B) = 0$ if and only if $\rho_{AB} = \rho_A \otimes \rho_B$. In other words, subsystems A and B do not contain any correlation between them. Indeed, the quantum mutual information is a measure for *total correlation* contained in a bipartite state in the following sense.

We remark that the total correlation is not equal to quantum entanglement. As the name suggested, the total correlation contains both classical correlation and quantum entanglement. A special case to check is that the total correlation of a separable state is not zero (because a separable state still contains classical correlation).

Given n copies of a bipartite state ρ_{AB}, we argue that the amount of noise (or randomness) Alice and Bob have to locally inject in order to bring it to the tensor product state $\rho_A \otimes \rho_B$:

$$\frac{1}{N} \sum_{i=0}^{N-1} (U_A^i \otimes V_B^i) \rho_{AB}^{\otimes n} (U_A^i \otimes V_B^i)^\dagger \approx_\varepsilon \rho_A^{\otimes n} \otimes \rho_B^{\otimes n}. \tag{85}$$

Theorem 18.

$$\lim_{n\to\infty} \frac{1}{n} \log N = I(A:B)_\rho. \tag{86}$$

4 Negativity of Conditional von Neumann Entropy and State Merging

The protocol of quantum state merging is as follows. Alice and Bob initially shared a state ρ_{AB}, and let $|\psi^\rho\rangle_{ABR}$ be the purification of ρ_{AB}. In addition, Alice and Bob also pre-shared a maximally entangled state $|\Phi_L\rangle_{T_AT_B}$ of dimension L:

$$|\Phi_L\rangle_{T_AT_B} = \frac{1}{\sqrt{L}}\sum_{i=1}^{L}|i\rangle_{T_A}\otimes|i\rangle_{T_B}.$$

Alice and Bob are allowed to perform LOCC, denoted by $\mathcal{E} : AT_ABT_B \to T'_AT'_BB\tilde{B}$, and let the state

$$\omega_{T'_AT'_BB\tilde{B}R} = \mathcal{E}(|\psi^\rho\rangle\langle\psi^\rho|^{\otimes n}\otimes|\Phi_L\rangle\langle\Phi_L|).$$

The protocol has ϵ-error if

$$\|\omega_{T'_AT'_BB\tilde{B}R} - |\Phi_K\rangle\langle\Phi_K|_{T'_AT'_B}\otimes|\psi^\rho\rangle\langle\psi^\rho|^{\otimes n}_{B\tilde{B}R}\|_1 \le \epsilon, \tag{87}$$

where $|\psi^\rho\rangle_{B\tilde{B}R}$ is same as $|\psi^\rho\rangle_{ABR}$ except that now the subsystem A has been transferred to Bob's side and denoted by $\tilde{B}$. Define the entanglement cost of the quantum state merging protocol to be

$$\frac{1}{n}(\log L - \log K) \tag{88}$$

Theorem 19 (Quantum State Merging)**.**

$$\lim_{n\to\infty}\frac{1}{n}(\log L - \log K) = H(A|B)_\rho. \tag{89}$$

The quantum state merging provides an operational interpretation to the conditional von Neumann entropy $H(A|B)_\rho$. When the $H(A|B)_\rho$ is negative, it means that the protocol generated entanglement instead of consuming it.

References

[1] Charles H. Bennett, Herbert J. Bernstein, Sandu Popescu, and Benjamin Schumacher, *Concentrating partial entanglement by local operations*, Phys. Rev. A **53** (1996Apr), 2046–2052.

[2] T. Cover and J. Thomas, *Wiley-interscience*, Cambridge University Press, 2006.

[3] E. Brian Davies and John T. Lewis, *An operational approach to quantum probability*, Comm. Math. Phys. **17** (1970), no. 3, 239–260.

[4] Patrick Hayden and Andreas Winter, *Communication cost of entanglement transformations*, Phys. Rev. A **67** (2003Jan), 012326.

[5] Hoi-Kwong Lo and Sandu Popescu, *Classical communication cost of entanglement manipulation: Is entanglement an interconvertible resource?*, Phys. Rev. Lett. **83** (1999Aug), 1459–1462.

[6] M. A. Nielsen, *Conditions for a class of entanglement transformations*, Phys. Rev. Lett. **83** (1999Jul), 436–439.

[7] Michael A. Nielsen and Isaac L. Chuang, *Quantum computation and quantum information*, Cambridge University Press, New York, 2000.

[8] Asher Peres and William K. Wootters, *Optimal detection of quantum information*, Phys. Rev. Lett. **66** (1991Mar), no. 9, 1119–1122.